ORIGINAL ARTICLE INFECTIOUS DISEASES

Predicting the risk of *Clostridium difficile* infection following an outpatient visit: development and external validation of a pragmatic, prognostic risk score

J. L. Kuntz[1], E. S. Johnson[1], M. A. Raebel[2,3], R. W. Platt[4], A. F. Petrik[1], X. Yang[1], M. L. Thorp[1,5], S. J. Spindel[6], N. Neil[7,8] and D. H. Smith[1]

1) *Kaiser Permanente Northwest, Center for Health Research, Portland, OR,* 2) *Kaiser Permanente Colorado, Institute for Health Research, Denver, CO,* 3) *Skaggs School of Pharmacy and Pharmaceutical Sciences, University of Colorado, Aurora, CO, USA,* 4) *McGill University Health Centre Research Institute, Department of Epidemiology and Biostatistics, Montréal, Québec, Canada,* 5) *Kaiser Permanente Northwest, Department of Nephrology,* 6) *Kaiser Permanente Northwest, Department of Infectious Diseases, Portland,* 7) *Decision Research, Eugene, OR and* 8) *University of Washington, Seattle, WA, USA*

Abstract

Increasing morbidity related to *Clostridium difficile* infection (CDI) has heightened interest in the identification of patients who would most benefit from recognition of risk and intervention. We sought to develop and validate a prognostic risk score to predict CDI risk for individual patients following an outpatient healthcare visit. We assembled a cohort of Kaiser Permanente Northwest (KPNW) patients with an index outpatient visit between 2005 and 2008, and identified CDI in the year following that visit. Applying Cox regression, we synthesized *a priori* predictors into a CDI risk score, which we validated among a Kaiser Permanente Colorado (KPCO) cohort. We calculated and plotted the observed 1-year CDI risk for each decile of predicted risk for both cohorts. Among 356 920 KPNW patients, 608 experienced CDI, giving a 1-year incidence of 2.2 CDIs per 1000 patients. The Cox model differentiated between patients who do and do not develop CDI: there was a *C*-statistic of 0.83 for KPNW. The simpler points-based risk score, derived from the Cox model, was validated successfully among 296 550 KPCO patients, with no decline in the area under the receiver operating characteristic curve: 0.785 (KPNW) vs. 0.790 (KPCO). The predicted risk for CDI agreed closely with the observed risk. Our CDI risk score utilized data collected during usual care to successfully identify patients who developed CDI, discriminating them from patients at the lowest risk for CDI. Our prognostic CDI risk score provides a decision-making tool for clinicians in the outpatient setting.

Keywords: *Clostridium difficile*, cohort study, model development, model validation, risk prediction models
Original Submission: 22 April 2014; **Revised Submission:** 5 October 2014; **Accepted:** 1 November 2014
Editor: M. Paul
Article published online: 11 November 2014

Corresponding author: J. L. Kuntz, The Center for Health Research, Kaiser Permanente Northwest, 3800 N. Interstate Avenue, Portland, OR 97227, USA
E-mail: jennifer.l.kuntz@kpchr.org

Introduction

Clostridium difficile infection (CDI) is recognized as a significant cause of morbidity and mortality [1]. Antimicrobial use, advanced age, underlying and severe comorbidity and prolonged hospitalization are well-known predictors of CDI [2–7]; however, physicians lack a tool that can quantify a patient's risk on the basis of demographic characteristics and medical history.

Although risk scores have commonly been developed to identify populations at high risk for chronic disease or its complications [8–11], prognostic risk modelling has been used less frequently for infectious diseases. Several scores have been created to identify hospitalized patients at risk for incident and recurrent CDI [12–17], but no risk prediction score has been

http://dx.doi.org/10.1016/j.cmi.2014.11.001

developed to identify risk for incident CDI among ambulatory patients.

Risk prediction in the ambulatory setting could allow clinicians to identify patients at the highest risk for CDI and modify individual risk factors, such as antimicrobial use, before these patients become ill. Patients at high risk for CDI could also be targeted for education about their CDI risk and recognition of symptoms. Our objective was to develop and validate a prognostic risk score that can be used to quantify the 1-year risk of CDI by using patient characteristics that are readily available during routine care delivery.

Study population and methods

Study design and patient selection

We conducted a cohort study among Kaiser Permanente Northwest (KPNW) members to develop a CDI risk score based on factors present during an outpatient visit that impact on the development of CDI. We then validated the risk score among a similarly defined cohort of Kaiser Permanente Colorado (KPCO) patients.

We selected patients with one or more outpatient primary or specialty care visits between 1 June 2005 and 30 September 2008, a time period during which both KPNW and KPCO used the same *C. difficile* testing approaches. For patients having multiple visits, we randomly selected one to serve as the index encounter. Study-eligible patients were aged 20–89 years, had at least 1 year of continuous health plan enrolment and prescription drug coverage prior to the index outpatient visit (i.e. index date), and had no evidence of CDI (e.g. diagnosis code, positive test result, or vancomycin dispensing) in the 180 days before and including the index date. Patients were allowed to have up to a 90-day gap in enrolment and still be considered to be 'continuously' enrolled. The same inclusion and exclusion criteria were used for the development and validation cohorts.

During the study time period, each of these health plans had ~450 000 members. Electronic databases provided information about enrolment, pharmacy dispensing, demographics, and clinical measures, including diagnoses, laboratory results, and healthcare utilization. The study was approved by Institutional Review Boards at both health plans.

Outcome measurement

We followed patients from their index date until the first observed CDI or for 12 months, whichever came first. We analysed time to first occurrence of CDI, defined as: (a) an International Classification of Diseases, 9th Revision (ICD-9) diagnosis code 008.45 for 'Intestinal infection due to *C. difficile*' from an outpatient encounter or hospitalization; or (b) a positive *C. difficile* toxin test in the outpatient setting. We also required positive toxin test results to be associated with (a) outpatient dispensing of metronidazole or vancomycin in the 7 days before or after the result date, or (b) an encounter for *C. difficile* (ICD-9 code 008.45) in the 7 days after the result date.

The CDI incident date was defined by the first indication of CDI (e.g. diagnosis code; date of vancomycin dispensing prior to a positive test). We did not follow patients after the first observed CDI. Patients were censored upon death or discontinuation of health plan enrolment. We also did not follow patients after 30 September 2008; thus, individuals with an index date after 30 September 2007 had <1 year of follow-up.

Predictors of CDI

We identified predictors of CDI from a review of the scientific literature [2–7]. We sought to develop a pragmatic risk score; thus, we considered the ease and reliability with which characteristics could be accessed from the electronic medical record (EMR) and the population prevalence of each characteristic. To avoid over-fitting to the development data, we limited the number of potential characteristics and their degrees of freedom [18].

Patient characteristics were obtained from coded information in the EMR. For time-varying characteristics—such as healthcare utilization, medication use, renal dialysis, and chemotherapeutic therapies—we collected data for the 60-day period before the index visit. The presence of comorbid conditions was assessed in the 1 year before the index date. Diagnosis codes and specific medications used to identify risk score characteristics are available upon request.

Development of the CDI risk score among the KPNW population

We used Cox regression to evaluate baseline characteristics that might predict CDI in the 1 year following an outpatient visit. Statistical significance did not guide the development of the model. Instead, we included all predictors identified *a priori*. We modelled categorical variables by using indicator variables. Continuous variables were modelled by use of a restricted cubic spline, which is one approach for describing a non-linear relationship between a predictor and an outcome [19]. The number and location of the knots for each restricted cubic spline were based on sample size, number of events, and available degrees of freedom relative to the expected explanatory power of each variable [20]. We included all *a priori* baseline characteristics, although we observed too few CDI events to evaluate interactions between characteristics.

We translated the Cox regression coefficients into a points-based risk score in which a higher number of points indicates a

higher CDI risk [11]. To accomplish this, the linear predictor in the Cox model was mapped to the corresponding 1-year CDI risk. The components of the linear predictor were then rescaled to an arbitrary axis in which a score of zero points was assigned to the lowest-risk category for each variable, with increasing points representing proportionate increases in the linear predictor. These risk score points approximate the hazard ratio for CDI. The points-based risk score offers decision-makers a practical tool for approximating the risk of CDI without the need to calculate the risk by using the full Cox regression equation and its coefficients.

Evaluation of risk score performance

We assessed predictive performance of the Cox model in the development cohort by examining discrimination, a measure of how well the model distinguishes who will and will not develop CDI, and calibration, a measure of how closely the predicted risk agrees with what was actually observed. We measured discrimination by the use of Harrell's *C*-statistic [21,22], the *D*-statistic, and the R^2-statistic. Harrell's *C*-statistic measures the proportion of all patient pairs in which predictions and outcomes are concordant. The *D*-statistic measures separation of survival curves. The R^2-statistic measures variation in CDI risk explained by the model. To show discrimination and calibration, we also plotted failure curves for observed and predicted CDI risk among KPNW patients according to deciles of predicted risk, as determined by the model [23].

External validation of the CDI risk score among the KPCO population

Because decision-makers in other settings are more likely to use a points-based risk score than a Cox regression equation, we validated CDI predictions in the KPCO population by using our points-based risk score. Using patient-level data from the

TABLE 1. Baseline characteristics for patients without *Clostridium difficile* infection (CDI) in the Kaiser Permanente Northwest (KPNW) development and Kaiser Permanente Colorado (KPCO) validation cohorts and the subgroup of patients in each cohort who experienced CDI during the 1 year after an outpatient visit

	KPNW		KPCO	
Patient characteristics	**Patients without CDI (*n* = 356 312)**	**Patients with CDI (*n* = 608)**	**Patients without CDI (*n* = 295 930)**	**Patients with CDI (*n* = 620)**
Age (years), mean (SD)	48.2 (16.5)	64.9 (16.9)	49.6 (17.2)	66.9 (16.5)
Age (years), *n* (%)				
20–29	54 859 (15.4)	22 (3.6)	42 718 (14.4)	17 (2.7)
30–39	63 652 (17.9)	36 (5.9)	51 225 (17.3)	35 (5.7)
40–49	70 472 (19.8)	52 (8.6)	56 996 (19.3)	48 (7.7)
50–59	77 359 (21.7)	94 (15.5)	59 652 (20.1)	85 (13.7)
60–69	50 303 (14.1)	128 (21.1)	41 042 (13.9)	103 (16.6)
70–79	25 118 (7.1)	126 (20.7)	30 315 (10.2)	162 (26.1)
80–89	14 549 (4.1)	150 (24.7)	13 982 (4.7)	170 (27.4)
Female, *n* (%)	195 947 (55.0)	353 (58.1)	167 094 (56.5)	377 (60.8)
Days of hospitalization in the previous 60 days				
0	346 150 (97.2)	493 (81.1)	285 769 (96.6)	505 (81.5)
1–6	8431 (2.4)	59 (9.7)	8786 (3.0)	69 (11.1)
≥7	1731 (0.5)	56 (9.2)	1375 (0.5)	46 (7.4)
Stay in a communal-living healthcare facility in the previous 60 days, *n* (%)	710 (0.2)	33 (5.4)	884 (0.3)	35 (5.7)
Chronic kidney disease, *n* (%)				
No diagnosis of chronic kidney disease and no history of dialysis	350 565 (98.4)	509 (83.7)	283 538 (95.8)	472 (76.1)
Chronic kidney disease diagnosis and no history of dialysis	4653 (1.3)	73 (12.0)	11 484(3.9)	111 (17.9)
History of dialysis	1094 (0.3)	26 (4.3)	908 (0.3)	37 (6.0)
Inflammatory bowel disease, *n* (%)	1976 (0.6)	16 (2.6)	1221 (0.4)	13 (2.1)
Immunosuppression in the previous 60 days, *n* (%)	13 415 (3.8)	107 (17.6)	11 602 (3.9)	128 (20.7)
Chronic pulmonary disease, *n* (%)	40 568 (11.4)	163 (26.8)	33 615 (11.4)	181 (29.2)
Rheumatological disease, *n* (%)	3422 (0.96)	26 (4.3)	3529 (1.2)	33 (5.3)
Diabetes, *n* (%)	31 801 (8.9)	165 (27.1)	23 566 (8.0)	139 (22.4)
Cardiovascular disease, *n* (%)	20 546 (5.8)	195 (32.1)	23 164 (7.9)	202 (32.6)
Liver disease, *n* (%)	969 (0.3)	14 (2.3)	762 (0.3)	14 (2.3)
Malignancy and metastatic solid tumour, *n* (%)	11 814 (3.3)	94 (15.5)	12 210 (4.1)	119 (19.2)
Chemotherapeutic procedures or therapies in the previous 60 days, *n* (%)	15 816 (4.4)	86 (14.1)	16 367 (5.5)	111 (17.9)
Use of β-lactam/β-lactamase inhibitors in the previous 60 days, *n* (%)	10 425 (2.9)	53 (8.7)	10 181 (3.4)	45 (7.3)
Use of cephalosporins in the previous 60 days, *n* (%)	7608 (2.1)	74 (12.2)	4361 (1.5)	30 (4.8)
Use of clindamycin in the previous 60 days, *n* (%)	2528 (0.7)	36 (5.9)	2084 (0.7)	17 (2.7)
Use of fluoroquinolones in the previous 60 days, *n* (%)	6049 (1.7)	65 (10.7)	5831 (2.0)	84 (13.6)
Use of macrolides in the previous 60 days, *n* (%)	5239 (1.5)	30 (4.9)	4286 (1.5)	20 (3.2)
Use of sulphonamides in the previous 60 days, *n* (%)	5788 (1.6)	27 (4.4)	3138 (1.1)	23 (3.7)
Use of tetracyclines in the previous 60 days, *n* (%)	5696 (1.6)	15 (2.5)	5070 (1.7)	17 (2.7)
Use of gastric acid suppressants in the previous 60 days, *n* (%)	32 454 (9.1)	161 (26.5)	27 491 (9.3)	136 (21.9)

SD, standard deviation.

TABLE 2. Adjusted hazard ratios and number of risk score points for patient characteristics included in the Cox regression model to predict *Clostridium difficile* infection during the 1 year after an outpatient visit among Kaiser Permanente Northwest members

Patient characteristics	Hazard ratio (95% CI)	Risk points
Age (years)[a]		
20–24	—	2
25–29	—	3
30–34	—	5
35–39	—	8
40–44	—	12
45—49	—	18
50–54	—	28
55–59	—	38
60–64	—	49
65–69	—	59
70–74	—	69
75–79	—	79
80–84	—	90
85–89	—	100
Sex		
Male	Reference	0
Female	1.16 (0.99–1.37)	6
Days of hospitalization in the previous 60 days, *n* (%)		
0	Reference	0
1–6	1.21 (0.89–1.64)	8
≥7	2.42 (1.67, 3.51)	37
Stay in a communal-living healthcare facility, *n* (%)		
No stay	Reference	0
History of a stay	1.78 (1.18–2.68)	24
Chronic kidney disease, *n* (%)		
No diagnosis of chronic kidney disease and no history of dialysis	Reference	0
Chronic kidney disease diagnosis and no history of dialysis	2.14 (1.62–2.84)	32
History of dialysis	1.83 (1.16–2.91)	25
IBD, *n* (%)		
No IBD	Reference	0
IBD	2.83 (1.71–4.67)	43
Immunosuppression, *n* (%)		
No immunosuppression	Reference	0
Immunosuppression	1.72 (1.34–2.20)	22
Chronic pulmonary disease, *n* (%)		
No chronic pulmonary disease	Reference	0
Chronic pulmonary disease	1.15 (0.95–1.4)	6
Rheumatological disease, *n* (%)		
No rheumatological disease	Reference	0
Rheumatological disease	1.13 (0.75–1.72)	5
Diabetes, *n* (%)		
No diabetes	Reference	0
Diabetes	1.36 (1.11–1.66)	13
Cardiovascular disease, *n* (%)		
No cardiovascular disease	Reference	0
Cardiovascular disease	1.57 (1.26–1.94)	19
Liver disease, *n* (%)		
No liver disease	Reference	0
Liver disease	3.13 (1.83–5.36)	47
Malignancy and metastatic solid tumour, *n* (%)		
No malignancy or metastatic solid tumour	Reference	0
Malignancy or metastatic tumour	1.72 (1.35–2.18)	22
Chemotherapeutic procedures or therapies, *n* (%)		
No chemotherapy	Reference	0
Chemotherapy	1.34 (1.05–1.72)	12
Use of β-lactam/β-lactamase inhibitors, *n* (%)		
No use	Reference	0
Any use	1.93 (1.45–2.58)	27
Use of cephalosporins, *n* (%)		
No use	Reference	0
Any use	2.51 (1.94–3.3)	38
Use of clindamycin, *n* (%)		
No use	Reference	0
Any use	4.1 (2.89–5.82)	58
Use of fluoroquinolones, *n* (%)		
No use	Reference	0
Any use	1.84 (1.39–2.43)	25
Use of macrolides, *n* (%)		
No use	Reference	0
Any use	1.66 (1.13–2.43)	21
Use of sulphonamides, *n* (%)		
No use	Reference	0
Any use	1.17 (0.79–1.74)	6
Use of tetracyclines, *n* (%)		
No use	Reference	3
Any use	0.94 (0.56–1.58)	0
Use of gastric acid suppressants, *n* (%)		
No use	Reference	0
Any use	1.34 (1.1–1.63)	12

IBD, inflammatory bowel disease.
[a]Age was modelled by the use of restricted cubic splines; the method does not give a summary hazard ratio and CI per unit increase.

KPCO cohort, we 'scored' each patient according to their baseline characteristics. After calculating each patient's risk score points, we were able to approximate their predicted 1-year risk for CDI. (Appendix 1 includes a summary table of the relationship between risk score points and the predicted 1-year risk; the complete version of the table is available upon request.)

Using KPCO data that included each patient's total risk score points and whether or not the patient actually developed CDI, we were able to fit a logistic regression model and calculate the area under the receiver operating characteristic (ROC) curve, which assesses discrimination—similar to the *C*-statistic. We repeated those steps for the KPNW cohort to calculate its area under the ROC curve, so that we could compare directly the performance for the points-based risk score (instead of the Cox regression equation and its exact coefficients).

We then divided the KPCO cohort into deciles of predicted risk. For each decile of predicted risk, we used patient-level data to calculate the Kaplan–Meier observed 1-year risk for CDI to assess calibration. We plotted those observed and predicted risks by decile for the KPCO cohort. Statisticians refer to this approach as a level 2 external validation of the Cox model, because it applies the risk score from a development population to a new population to assess discrimination and calibration [24].

Results

Development of the risk score among the KPNW population

Among 356 920 KPNW patients with an index outpatient visit, we observed 608 CDIs, giving a 1-year incidence of 2.2 CDIs per 1000 patients (95% CI 2.0–2.4). Patients who developed CDI were older and more likely to have a history of hospitalization or a stay in a communal-living healthcare facility (Table 1). Patients who developed CDI were also more likely to have recently received antibiotics and gastric acid suppressants or to have a history of immunosuppression or chemotherapy. All comorbid conditions occurred more frequently among patients who developed CDI.

The patient characteristics that contributed >30 points to the risk score, indicating an approximate doubling of risk, were: age of ≥55 years (38–100 points, depending on age category); hospitalization of ≥7 days (37 points); liver disease (47 points); inflammatory bowel disease (43 points); and cephalosporin use (38 points) or clindamycin use (58 points) (Table 2). The 22 characteristics in the score accounted for 27 degrees of freedom. The slope-shrinkage statistic, which assesses the extent to which the model may be over-fitted to the development data, was 0.99 (on a scale from 0 to 1, where 1 indicates no over-fitting).

Summary statistics from the internal validation of the prediction model indicated effective performance in the KPNW cohort. Patients in the highest-risk decile were nearly 60 times more likely to suffer a CDI (1-year incidence of 9.8 CDIs/1000 patients) than patients in the lowest-risk decile (0.2 CDIs/1000 patients) (Fig. 1). The bootstrap-corrected *C*-statistic (0.83) and the *D*-statistic (2.35) indicate effective discrimination or separation of the high-risk and low-risk patients. The R^2-statistic (56.82) indicates that the prediction model explained 56.82% of the variation in CDI risk. The predicted risk was closely aligned with observed events for the highest-risk patients. Agreement (calibration) was adequate for the remaining deciles (Fig. 1). The simpler points-based risk scoring system retained its effective discrimination, although it was less effective than the prediction model based on the Cox regression equation with exact coefficients (area under the ROC curve of 0.786).

External validation of risk score among the KPCO population

We validated the score among 296 550 similarly-defined KPCO patients; 620 of these patients experienced CDI, yielding a 1-year incidence of 2.3 CDIs per 1000 patients (95% CI 2.1–2.5). Table 1 shows the baseline characteristics for KPCO patients. The simpler, points-based risk score (developed in the KPNW cohort) performed comparably in the KPCO cohort in terms of discrimination, as shown by the area under the ROC

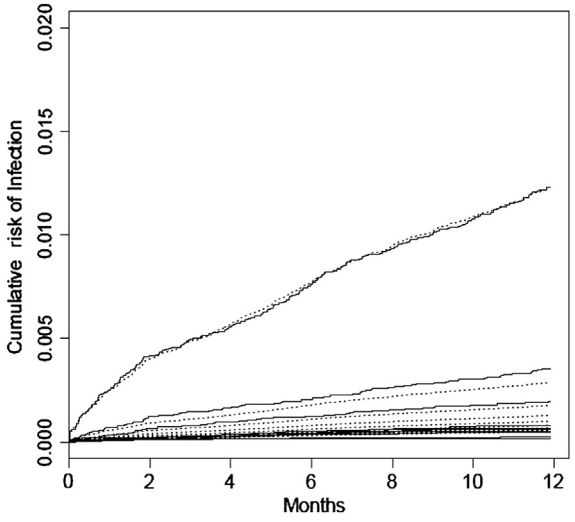

FIG 1. Failure curves showing the observed risk (solid lines) and the predicted risk (dotted lines) of *Clostridium difficile* infection during the first year after an outpatient visit among Kaiser Permanente Northwest patients according to deciles of predicted risk, as determined by the risk score.

curve (0.790). Fig. 2 demonstrates successful calibration, or agreement, of the points-based risk score for each decile of predicted risk.

Discussion

We developed and validated a risk score that successfully discriminated between patients at the highest and lowest risk for CDI. Our predictions agreed closely with the observed risk for CDI. The risk score provides important information about CDI risk for patients who are most likely to benefit from clinician recognition of this risk, and for specific clinical decision-making, clinical monitoring, and/or education. For example, in the KPNW population, using our risk score could help clinicians to focus on the 10% of patients who will experience 57% of the CDIs.

Our risk score is the first to be developed and externally validated to predict incident infection in an unselected, outpatient routine care setting. This reliable and accurate tool will enable clinicians to use combinations of predictors to estimate the probability that CDI will occur, and to identify patients in whom recognition of elevated risk can prompt risk management efforts (e.g. more judicious antimicrobial use). In the future, this score might be used to select high-risk patients for clinical trials, or in turn, to determine patients for whom primary prevention (i.e. vaccination) would be warranted.

For example, on the basis of this model, practitioners might choose a threshold for targeting patients for increased risk management. To illustrate: if a physician chose the 90th percentile of the risk score (i.e. the top decile) as a threshold, he or she would focus attention on patients who experienced 57.2% of CDIs (95% CI 53.2–61.2%). The specificity for that threshold is 90.3% (95% CI 90.2–90.4%). One per cent of patients in the top decile developed CDI (positive predictive value of 1%; 95% CI 0.9–1.11%). The negative predictive value was 99.9% (95% CI 99.9–99.9%). In order to score in this top decile, a patient would require at least 90 risk points. As an example, a 65-year-old patient with diabetes and recent fluoroquinolone use would score 97 points (age 65 years (59 points) + history of diabetes (13 points) + fluoroquinolone use (25 points) = 97 points).

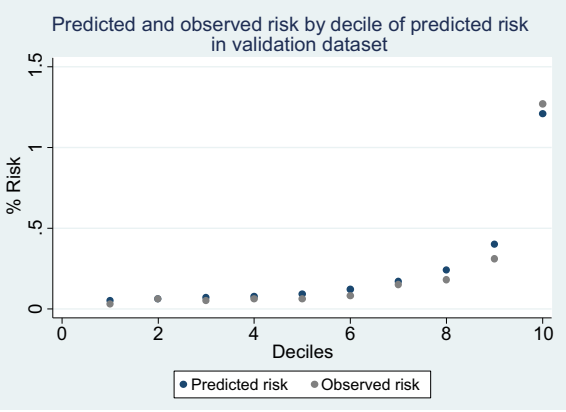

FIG 2. Observed and predicted risk of *Clostridium difficile* infection during the first year after an outpatient visit among Kaiser Permanente Colorado patients, according to deciles of predicted risk as determined by application of the points-based risk score. The figure demonstrates successful calibration for each decile of predicted risk. For instance, in the top decile of predicted risk, the predicted risk was slightly lower than the observed risk.

We developed our risk score to be pragmatic, meaning that this score can be easily calculated from readily available clinical and non-clinical information to inform decision-making at the point of care. In usual care, clinicians assess risk on the basis of known risk factors, although the complexity of such calculations can often be overwhelming, requiring physicians to consider individual factors without the support of a statistical algorithm to account for relationships between individual characteristics [25,26]. In contrast, our risk score is designed to work with data that are routinely collected and available, allowing for a patient's CDI risk to be automatically calculated and made directly available to clinicians via an EMR.

Our study has limitations. First, we did not include patients with a recent history of CDI, so the risk score is not tailored to estimate the risk of CDI recurrence. Second, we did not measure antimicrobial or other medication use during hospitalization. Next, because information about the date of microbiological testing and treatments dispensed were unavailable for hospitalized patients, we limited our identification of CDI in the inpatient setting to the use of diagnosis codes. Finally, it should be noted that the accuracy of the risk score's predictions and the low incidence of CDI mean that the positive predictive value is low; that is, most patients designated as 'high risk' will not go on to develop CDI.

Our risk score successfully discriminates between patients at the highest and lowest risk for CDI among patients seeking outpatient care. Clinicians who wish to use this risk score may want to validate its predictions in their populations and healthcare settings. Future efforts should focus on determining the impact of the risk score on patient management and outcomes, as compared with usual care [25].

Transparency declaration

Financial support for this study was provided by a contract with Sanofi Pasteur. The funding agreement ensured the authors'

independence in designing the study, interpreting the data, and writing and publishing the report.

Appendix 1. Risk score point ranges and corresponding ranges of predicted probabilities for *Clostridium difficile* infection in the 1 year following an outpatient visit

Risk score points	Predicted risk (%)
0–128	<1.00
129–157	1.00–1.99
158–174	2.00–2.99
175–186	3.00–3.99
187–196	4.00–4.99
197–203	5.00–5.99
204–210	6.00–6.99
211–216	7.00–7.99
217–221	8.00–8.99
222–225	9.00–9.99
≥226	≥10.00

References

[1] Elixhauser A, Jhung M. *Clostridium difficile*-associated disease in U.S. hospitals, 1993–2005: Statistical Brief #50. Rockville, MD: Agency for Health Care Policy and Research (US); February 2006.

[2] Gerding DN, Johnson S, Peterson LR, Mulligan ME, Silva Jr J. *Clostridium difficile*-associated diarrhea and colitis. Infect Control Hosp Epidemiol 1995;16(8):459–77.

[3] Bignardi GE. Risk factors for *Clostridium difficile* infection. J Hosp Infect 1998;40(1):1–15.

[4] Blondeau JM. What have we learned about antimicrobial use and the risks for *Clostridium difficile*-associated diarrhoea? J Antimicrob Chemother 2009;63(2):238–42.

[5] Dubberke ER, Reske KA, Yan Y, Olsen MA, McDonald LC, Fraser VJ. *Clostridium difficile*-associated disease in a setting of endemicity: identification of novel risk factors. Clin Infect Dis 2007;45(12):1543–9.

[6] Linsky A, Gupta K, Lawler EV, Fonda JR, Hermos JA. Proton pump inhibitors and risk for recurrent *Clostridium difficile* infection. Arch Intern Med 2010 May 10;170(9):772–8.

[7] Howell MD, Novack V, Grgurich P, Soulliard D, Novack L, Pencina M, et al. Iatrogenic gastric acid suppression and the risk of nosocomial *Clostridium difficile* infection. Arch Intern Med 2010 May 10;170(9):784–90.

[8] D'Agostino Sr RB, Vasan RS, Pencina MJ, Wolf PA, Cobain M, Massaro JM, et al. General cardiovascular risk profile for use in primary care: the Framingham Heart Study. Circulation 2008 Feb 12;117(6):743–53.

[9] Hippisley-Cox J, Coupland C. Individualising the risks of statins in men and women in England and Wales: population-based cohort study. Heart 2010 Jun;96(12):939–47.

[10] Apgar V. The newborn (Apgar) scoring system. Reflections and advice. Pediatr Clin North Am 1966 Aug;13(3):645–50.

[11] Sullivan LM, Massaro JM, D'Agostino Sr RB. Presentation of multivariate data for clinical use: The Framingham Study risk score functions. Stat Med 2004 May 30;23(10):1631–60.

[12] Garey KW, Dao-Tran TK, Jiang ZD, Price MP, Gentry LO, DuPont HL. A clinical risk index for *Clostridium difficile* infection in hospitalised patients receiving broad-spectrum antibiotics. J Hosp Infect 2008 Oct;70(2):142–7.

[13] Tanner J, Khan D, Timmons S. Using the Waterlow tool to predict *Clostridium difficile* infection risk in hospital settings. Nurs Times 2010 Sep 14;106(36):20–2.

[14] Tanner J, Khan D, Anthony D, Paton J. Waterlow score to predict patients at risk of developing *Clostridium difficile*-associated disease. J Hosp Infect 2009 Mar;71(3):239–44.

[15] Dubberke ER, Yan Y, Reske KA, Butler AM, Doherty J, Pham V, et al. Development and validation of a *Clostridium difficile* infection risk prediction model. Infect Control Hosp Epidemiol 2011 Apr;32(4):360–6.

[16] D'Agostino Sr RB, Collins SH, Pencina KM, Kean Y, Gorbach S. Risk estimation for recurrent *Clostridium difficile* infection based on clinical factors. Clin Infect Dis 2014 May;58(10):1386–93.

[17] Zilberberg MD, Reske K, Olsen M, Yan Y, Dubberke ER. Development and validation of a recurrent *Clostridium difficile* risk-prediction model. J Hosp Med 2014 Jul;9(7):418–23.

[18] Smith LR, Harrell FE, Mahlbaier LH. Problems and potentials in modeling survival. Rockville MD. US: Department of Health and Human Services: Medical Effectiveness Research Data Methods (Summary Report); 1992.

[19] Harrell Jr FE, Lee KL, Pollock BG. Regression models in clinical studies: determining relationships between predictors and response. J Natl Cancer Inst 1988 Oct 5;80(15):1198–202.

[20] Harrell FE. General aspects of fitting regression models. Regression modeling strategies. New York: Springer; 2001. p. 11–37.

[21] Harrell Jr FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. Stat Med 1996 Feb 28;15(4):361–87.

[22] Pencina MJ, D'Agostino Sr RB, D'Agostino Jr RB, Vasan RS. Evaluating the added predictive ability of a new marker: from area under the ROC curve to reclassification and beyond. Stat Med 2008 Jan 30;27(2):157–72.

[23] Lumley T, Kronmal RA, Cushman M, Manolio TA, Goldstein S. A stroke prediction score in the elderly: validation and Web-based application. J Clin Epidemiol 2002 Feb;55(2):129–36.

[24] Royston P, Altman DG. External validation of a Cox prognostic model: principles and methods. BMC Med Res Methodol 2013;13:33.

[25] Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research: application and impact of prognostic models in clinical practice. BMJ 2009 Jun 4;338:1373–7.

[26] Maddirala S, Khan A, Vincent A, Lau K. Effect of angiotensin converting enzyme inhibitors and angiotensin receptor blockers on serum potassium levels and renal function in ambulatory outpatients: risk factors analysis. Am J Med Sci 2008 Oct;336(4):330–5.